

REMARKS

Claims 1-12 are all the claims pending in the application.

As preliminary matter, an abstract of the disclosure as required by 37 CFR 1.72(b) is submitted on a separate sheet as required. Applicant asks the Examiner to acknowledge the claim of priority under 35 U.S.C. § 119 and to acknowledge receipt of certified copies of all of the priority documents. Further, the Examiner did not indicate whether the drawings are accepted or objected to. Applicant requests that the Examiner make this indication.

Claims 1-12 are rejected under 35 U.S.C. § 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. In particular, the Examiner states that in claims 1-12 the phrase “distinct” is indefinite and needs to be revised or deleted. The Examiner states that in claim 5 the phrase “product” is indefinite.

Applicant submits that the phrase “distinct” would be clear to a hypothetical person possessing the ordinary skill level in the pertinent art, data tabulation processing. The Cambridge Dictionary of American English defines the adjective “distinct” as different or not alike. The Harper Collins Dictionary of Mathematics defines the adjective “distinct” as not numerically identical. It is clear that “distinct” means only the different data values for a field, which takes notice of duplicate data values in a field.

Applicant submits that the use of the phrase “distinct” is necessary to claim the subject matter of the invention. The data tabulation involves processing a data file having a plurality of fields to count the number of records for each combination of distinct field values for the

multiple fields. Tabulation thus requires the computation of the number of records with a certain combination of distinct values for the specified fields. *See Specification Table 1; Table 2; Page 3, lines 24-30.* As seen in Table 2, from the data tabulation it is possible to easily identify the number of data records from the data file which have a particular combination of attributes of each data field. If the word “distinct” were deleted the method of the present invention would no longer be claimed.

Thus, the phrase is definite and is necessary to particularly point out the invention.

Applicant submits that the phrase “product” would be clear to a hypothetical person possessing the ordinary skill level in the pertinent art, data tabulation processing. The meaning of “product” is apparent from the specification or the accepted meaning of the phrase. The specification states that the number of numerical identifiers for each field corresponds to the number of distinct data values in the field. The number of cells in the resultant array is computed by taking the product of the number of distinct data values for the specified fields. For example, if fields P and Q contain D_P and D_Q distinct values, respectively, the number of cells in the result array is $D_P \times D_Q$. *See Specification Page 5, 17-22; Page 6, lines 20-26.* In addition, a person in the art would know that “product” is synonymous with multiplication.

Thus, the phrase is definite and is necessary to particularly point out the invention.

Claims 1-12 are rejected under 35 U.S.C. § 103(a) as being unpatentable over Sampson et al. (U.S. 5,212,639) in view of Rees et al. (U.S. 5,748,878). The Examiner sets forth a detailed statement of this rejection on pages 2-4.

Applicant submits that claim 1 as amended, which is a combination of claim 1 and a part of claim 2, is patentable over the cited references.

In claim 1 as amended, a pre-processing stage in which, for each individual field, each distinct data value is identified and allocated a numerical identifier unique for that field, and a mapping file for the numerical identifiers is generated. As discussed in the Specification on page 5 at lines 17-26, all individual field files 12 having discrete fields are submitted to a distinct code mapping unit 14. *See Fig. 1.* The distinct code mapping unit 14 reads through the individual field files 12, and for each individual field file assigns an incremental numerical code to all field values found within. This numerical code starts from 0 and increments by 1 with each distinct field value found therein. The method of the claimed invention does not take notice of the nature or quantity of the individual field values. The numerical identifier is incrementally assigned to distinct values as they are encountered. A mapping file 18 is produced for each field to capture the code assigned for each field value. In contrast, none of the references teach allocating a unique numerical identifier for each distinct data value without notice to nature or quantity of the individual field values, wherein a mapping file is generated.

Applicant submits that Sampson does not disclose or suggest the claimed feature. The Examiner argues that a preprocessing stage in which each data value is identified and allocated a numerical identifier unique for that field is inherent from classifying large volumes of raw data. However, Applicant submits that the Examiner is misinterpreting and/or misapplying the teachings of Sampson.

Sampson discloses that a list of items, all item numbers and associated descriptions that may be found in the parental set is utilized to generate item indicators. *See Col. 3, lines 28-30.* However, the assignments are based on the nature and quantity of the data entries. For example, one item indicator can represent a positive quantity of an item number and another item indicator can represent a negative quantity of the same item number. *See Col. 3, lines 31-34.* Furthermore, a list of all account numbers and associated descriptions, a chart of accounts, is utilized to generate account sections (item indicators). A mapping function is applied to the journal entry lines in the parental set to assign the appropriate account-sections resulting in an expanded parental set. *See Col. 3, lines 51-61.*

However, in claim 1 as amended, no knowledge of the individual data values contained in the field is required. The mapping file is generated as a result of the assignment of the indicator numbers, the indicator numbers are not assigned as a result of an already existing mapping file (chart of accounts).

Therefore, Sampson does not disclose or suggest allocating a unique numerical identifier for each distinct data value for each individual field, wherein a mapping file is generated.

Finally, the Examiner does not rely on Rees to suggest the feature above, and there is no teaching in the reference of the above claimed feature.

Since claims 2-12 depend on claim 1, Applicants submit that they are patentable at least by virtue of their dependency.

Amendment Under 37 C.F.R. § 1.111
U.S. Application No.: 09/582,716

Attorney Docket No.: Q58912

In view of the above, reconsideration and allowance of this application are now believed to be in order, and such actions are hereby solicited. If any points remain in issue which the Examiner feels may be best resolved through a personal or telephone interview, the Examiner is kindly requested to contact the undersigned at the telephone number listed below.

The USPTO is directed and authorized to charge all required fees, except for the Issue Fee and the Publication Fee, to Deposit Account No. 19-4880. Please also credit any overpayments to said Deposit Account.

Respectfully submitted,

A handwritten signature in black ink, appearing to be 'Richard C. Turner', written over a horizontal line.

Richard C. Turner
Registration No. 29,710

SUGHRUE MION, PLLC
2100 Pennsylvania Avenue, N.W.
Washington, D.C. 20037-3213
Telephone: (202) 293-7060
Facsimile: (202) 293-7860

Date: **December 19, 2002**

APPENDIX
VERSION WITH MARKINGS TO SHOW CHANGES MADE

IN THE CLAIMS:

The claims are amended as follows:

1. (Amended) A method for data tabulation processing of a data file having a plurality of records in a plurality of data fields, comprising:
 - i) a pre-processing stage in which, for each individual data field, each distinct data value is identified and allocated a numerical identifier unique for that field, wherein the preprocessing stage includes generating from said data file a mapping file which stores a correspondence between each of the distinct data values in the fields and the corresponding numerical identifiers; and
 - ii) a tabulation stage in which, for each data record a cell of a result array is determined based on the numerical identifiers for that record, and the result array cell incremented.
2. (Amended) A method as in claim 1, wherein the preprocessing stage includes generating from said data file an encoded data file containing the numerical identifiers for the data values in each field[, and a mapping file which stores a correspondence between each of the distinct data values in the fields and the corresponding numerical identifiers].

ABSTRACT

A system and method for high speed, high volume tabulation of data using a personal computer. The method involves two sub-processes, pre-processing and tabulation. In pre-processing the data field is separated into individual files, and each distinct value in each of the discrete fields is encoded with an incremental numerical code unique for that field. Tabulation of fields involves the creation of a one-dimensional result array in which each cell will contain the result of tabulation of each combination of distinct field values. The encoded files for all requested fields is then read and the relevant cell for each record is incremented, the relevant cell for each record being identified using a mapping algorithm based on the codes of the fields. This method and system allows tabulation to be done at high speed.